DTU Library

Countermeasures and Barriers

Petersen, Johannes

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Petersen, J. (2005). *Countermeasures and Barriers*. Nordic Nuclear Safety Research. NKS No. 113

Nordisk kernesikkerhedsforskning
Norrænar kjarnöryggisrannsóknir
Pohjoismainen ydinturvallisuustutkimus
Nordisk kjernesikkerhetsforskning
Nordisk kärnsäkerhetsforskning
Nordic nuclear safety research

NKS-113
ISBN 87-7893-174-6

Countermeasures and Barriers

Johannes Petersen
Technical University of Denmark

October 2005

Abstract

In 1973 Haddon proposed ten strategies for reducing and avoiding damages based on a model of potential harmful energy transfer (Haddon, 1973). The strategies apply to a large variety of unwanted phenomena. Haddon's pioneering work on countermeasures has had a major influence on later thinking about safety. Considering its impact it is remarkable that the literature offers almost no discussions related to the theoretical foundations of Haddon's countermeasure strategies. The present report addresses a number of theoretical issues related to Haddon's countermeasure strategies, which are:

- A reformulation and formalization of Haddon's countermeasure strategies
- An identification and description of some of the problems associated with the term "barrier"
- Suggestions for a more precise terminology based on the causal structure of countermeasures
- Extending the scope of countermeasures to include sign-based counter-measures

Key words

Countermeasures, barriers, Haddon's strategy, theoretical foundation

NKS-113
ISBN 87-7893-174-6

Electronic report, October 2005

The report can be obtained from
NKS Secretariat
NKS-775
P.O. Box 49
DK - 4000 Roskilde, Denmark

Phone +45 4677 4045
Fax +45 4677 4046
www.nks.org
e-mail nks@nks.org

Countermeasures and Barriers

NKS-R-07, Barriers, Control and Management

Johannes Petersen
Ørsted DTU, Automation
Technical University of Denmark
DK-2800 Kongens Lyngby
jop@oersted.dtu.dk

October 7, 2005

Introduction

In 1973 Haddon proposed ten strategies for reducing and avoiding damages based on a model of potential harmful energy transfer (Haddon, 1973). The strategies apply to a large variety of unwanted phenomena. Haddon's pioneering work on countermeasures has had a major influence on later thinking about safety. Considering its impact it is remarkable that the literature offers almost no discussions related to the theoretical foundations of Haddon's countermeasure strategies. The present report addresses a number of theoretical issues related to Haddon's countermeasure strategies, which are:

- A reformulation and formalization of Haddon's countermeasure strategies
- An identification and description of some of the problems associated with the term "barrier"
- Suggestions for a more precise terminology based on the causal structure of countermeasures
- Extending the scope of countermeasures to include sign-based countermeasures

In the report a means-end view of countermeasures is adopted. A countermeasure is an action that serves as a means of opposing some unwanted phenomena, whereas the entities participating in a countermeasure are considered as means of the countermeasure.

A formalization of Haddon's countermeasure strategies is proposed based on a model of causation that generalizes Haddon's energy transfer model (representing some unwanted phenomena). The chosen model gives primacy to causal action and the role of the underlying entities participating in action (e.g. agent and patient). By means of adopting a negative perspective on the determinants of causal action a set of generic countermeasures is derived. Preliminary results can be found in (Lind and Petersen, 2003).

The proposed formalization of Haddon's strategies throws new light on the theoretical foundation of countermeasure strategies. Furthermore, it widens the field of application to areas where the countermeasures based on an energy transfer model require too much forcing. The latter has been the primary motivation for the generalization of the MORT barrier analysis (see e.g. (Kingston, 2002) and (Kingston et al., 2004)).

The term "barrier" is widely used for the analysis and assessment in safety critical domains, such as nuclear power plants (e.g. (Svenson, 1991)). Despite its widespread use, the term "barrier" is somewhat problematic and in the safety literature its specific meaning lacks consensus. The present report outlines some of the problems associated with the term "barrier" and attempts to clarify its use based on an analysis of the causal structure of countermeasures, i.e. the relationship between a specific countermeasure and the participating entities.

In the model of causation the causal agent has no decision-making capabilities. Clearly this is a limitation if we want to consider the countermeasures for actions performed by humans. Accordingly a revised model is proposed where the causal agent is substituted by an intelligent agent.

In order to account for countermeasures of unwanted actions performed by intelligent agents it is necessary to extend the scope of countermeasures from physical actions to *communicative actions*, i.e. actions based on *signs* produced in order to influence the action of the agent interpreting them. The producer and the interpreter of the sign may be separated in space and time. A procedure, for example, is a sign produced in order to make an intelligent agent act in a specific way.

The present report discusses the means-end levels of communicative action based on Speech Act Theory (Austin, 1962)(Searle, 1969). A set of sign-based countermeasure types, that complements the causal countermeasure types, is proposed.

Haddon's countermeasure strategies

Haddon´s ten generic countermeasure strategies are outlined in Table 1. The strategies represent a generalization of accident countermeasures based on experience from a large variety of domains.

Haddon's Countermeasure Strategies
1. To prevent the initial marshalling of the form of energy
2. To reduce the amount of energy marshalled
3. To prevent the release of the energy
4. To modify the rate of spatial distribution of release of energy from its source
5. To separate in space or time the energy being released from the susceptible structure
6. To separate the energy being released from the susceptible structure by interposition of a material barrier
7. To modify the contact surface, subsurface, or basic structure which can be impacted
8. To strengthen the structure which might be damaged by the energy transfer
9. To move rapidly in detection and evaluation of damage and to counter its continuation and extension
10. All those measures which fall between the emergency period following the damaging energy exchange and the final stabilization of the process

Table 1. Haddon's countermeasure strategies (Haddon, 1973).

According to Haddon the sequence of the countermeasure strategies is *logical*. Presumably, this refers to the fact that the strategies are ordered according to the downstream energy flow path from the source to the target. A successful application of a countermeasure arrests the energy flow. Interestingly, Haddon argues that there is no logical reason why the rank order of loss-reduction countermeasures generally considered must parallel the sequence, or rank order, of causes contributing to the result of damaged people or property. As an example Haddon mentions the possibility of eliminating the damage of teacups by packaging them properly (the sixth strategy) even though they are placed in motion in the hands of the postal service.

Haddon mentions a single heuristics: The larger the amounts of energy involved in relation to the resistance to damage of structures at risk, the earlier the strategy must lie in the countermeasure sequence.

A model-based approach to countermeasures

Haddon has demonstrated that experiences of countermeasures across different domains can be generalized on the basis of an energy transfer model of unwanted phenomena.

A model of the unwanted phenomena is fundamental in order to specify different types of countermeasures. Furthermore, it enables a rigorous application of countermeasures, which facilitates a certain level of reproducibility. Nevertheless, approaches to countermeasure/barrier analysis that lack a model of the underlying mechanism of the unwanted phenomena are still quite common. Typically, such approaches define a countermeasure/barrier as something that hinders one event from causing another (see e.g. (Svenson (1991)).

Perceiving causality as a relation between events is a heritage from Hume's regularity theory of causality (Hume, 2000). This theory denies the common idea that causal action, by which an active cause produces an effect, can be experienced. So the causal relation reflects no more than a *cause as experienced* and an *effect as experienced*, and the fact of their co-appearance (regularity). The fact that we experience causal production is judged by Hume to be illusory and have a psychological explanation.

Based on an event view of causation it is not possible to discriminate systematically between different kinds of countermeasures. Furthermore, the specification of countermeasures/barriers will depend on the identification of events, which is typically strongly subjective (Rasmussen, 1991).

Formalizing Haddon's countermeasure strategies

By substituting the underlying energy transfer model with a model of causation this section describes a formalization of Haddon's countermeasure strategies described in Table 1. Other authors have attempted to formalize countermeasure strategies. For instance, Kingston (2002) has suggested to generalize the energy transfer model (the model on which the MORT barrier analysis is founded) in order to extend the field of application of the MORT program. The problem is that, in some domains, the energy approach requires too much forcing (Kingston et al., 2004). By means of a model based on the concept of agency Kingston (2002) proposed a generalization of the energy transfer model behind the barrier analysis in the MORT program. The model proposed in this report also incorporates the notion of agency, but originates mainly from a realist approach to causation described by Harré and Madden (1975). This model provides an alternative to the regularity theory of causality and gives primacy to causal action and the participating entities. Events only have a secondary role to play in this model.

A model of causation

Instead of perceiving causality as a relation among events the present model describes causality in terms of a process of power-actualization. That is, the focus is not on events but on the underlying entities that interact when their powers (and liabilities) are actualized.

The model makes a distinction between *intrinsic* and *extrinsic* determinants of causal action. The intrinsic determinants refer to the entities having powers, which they possess in virtue of their intrinsic natures (Harré and Madden, 1975). Powers represent potentialities for causal action that are actualized when appropriate conditions are in place. In the present model events are reduced to extrinsic determinants. That is, events can be identified as having a role in the power-actualization only if they can be shown to trigger the action of a powerful entity.

The notion of power is inherently active and reflects the property of an entity (a *causal agent*) with the potentiality to produce a change or alteration in some other entity. The notion of liability is the passive counterpart of power and reflects the property of an entity (a *patient*) with the potentiality to undergo a change or alteration under the influence of some other entity. Because the same entity can have powers as well as liabilities it can have different roles in relation to causal action.

As opposed to an intelligent agent a causal agent has no decision-making capabilities. A causal agent is an entity with the power to perform some action that is not intentional, at least from the part of the causal agent. The relationship between causal action and the powerful entities participating in action is illustrated in Figure 1. The entity X has an agent role whereas Y has a patient role in the causal action.

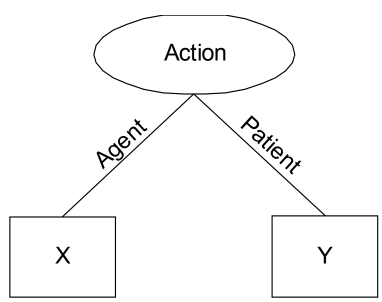

Figure 1. The structure of causal action.

This account of causation is similar in structure to the linguistic account of the *semantic deep structure* of a sentence. According to the theory of semantic deep structure in linguistics, introduced by Fillmore (1968), the semantic content of a sentence S can be represented as a proposition P, specified in terms of relationships between a verb V and the other constituents of the sentence (typically noun phrases). The main point is that such relationships can be covered by a limited number of fixed *cases* or *semantic roles*, such as *agent*, *object*, *instrument,* etc. A predicate (the verb) and set of arguments (the semantic roles) define the propositions used to represent the semantic deep structure of a sentence. For instance, the following two sentences: 1) *"John opened the door with a key"* and 2) *"John used a key to open the door"* have the same semantic deep structure which can be represented by the proposition: (open (Agent John) (Object door) (Instrument key)).

The similarity between the causal model and the linguistic account of natural language semantics means that AI modeling schemes such as *semantic nets* (Brachman, 1979) and *conceptual graphs* (Sowa, 1984) can be used as tools for representations based on the causal model described above.

In the remaining part of the report the structure of action, defined by the roles describing the relationship between an action and its participating entities, is also referred to as the semantic structure of action.

Action Possibilities

An action is actualized when the powers and liabilities of the entities participating in the action are actualized. A precondition for actuality is that the action is *possible* in the first place. In turn, action possibility depends on *potentiality* and *opportunity* for action.

The *action potentiality* is determined by the *powers* and *liabilities* of the entities (as described above). An action is potential when the causal agent has the power to do A and the patient has the liability to undergo A. The *action opportunity* exists when the spatio-temporal settings allow the agent to act on the patient (Lind, 2000). In other words, the patient should be in the "reach" of the agent. Apart from the action possibility, the actualization of an action typically requires certain triggering conditions. Consequently, the above can be summarized roughly by the following sentence:

Potentiality & Opportunity & Triggering conditions = Actuality

Potentiality represents intrinsic determinants of causal action while opportunity and triggering conditions represent extrinsic determinants. Below a more detailed description is given. See also Figure 2.

An action A is *possible* only if:

1) The agent and the patient have the *potentiality* to enter into action.
    a. There is an agent with the *power* to do A,
    b. A patient with the *liability* to undergo A, and

2) The agent and the patient have the *opportunity* to take part in the action A (the agent can 'reach' the patient).

An action A is *actualized* only if:
1) Action A is possible,
2) The power of the agent to do A is actualized by triggering conditions and
3) The action A is not counteracted.

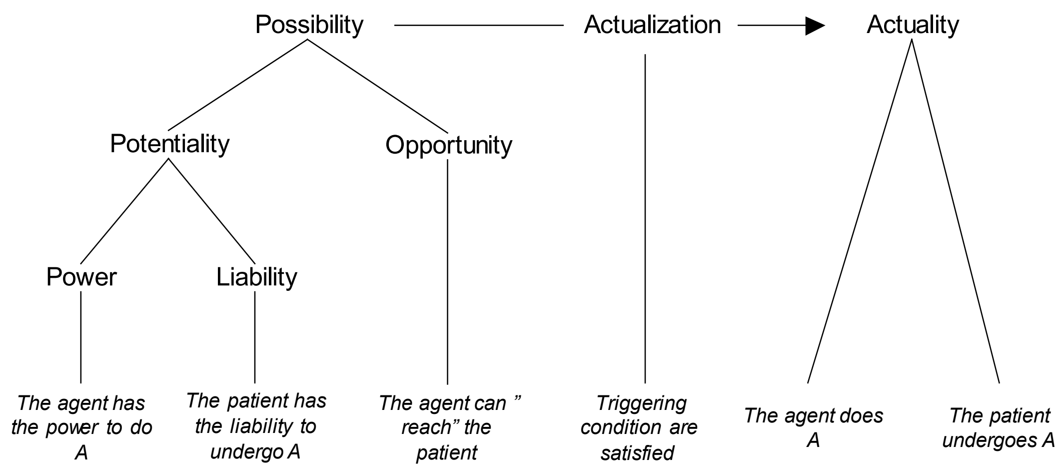

Figure 2. Overview of the determinants of causal action.

Using the model of causation to formalize Haddon's countermeasure strategies

In the above the causal determinants (intrinsic and extrinsic) were defined for causal actions in general. By means of adopting a negative perspective on the determinants of an unwanted action it is possible to derive a formalized account of countermeasure types, i.e. the ways in which an unwanted causal action can be opposed (i.e. prevented or mitigated). Note that we are referring to two types of action here: 1) an action that is unwanted, and 2) an action which is supposed to counteract the unwanted action. Figure 3 illustrates the means-end relationship between the two action types (A and B). More specifically, action B is a means of counteracting action A and we refer to Action B as a countermeasure.

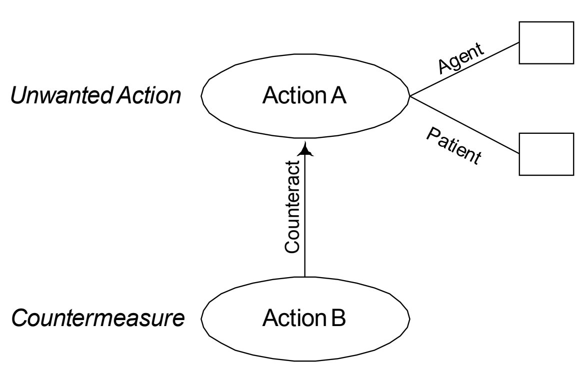

Figure 3. Means-end relation between an unwanted action A and a countermeasure B.

		Formalized Countermeasure Types	Haddon's Countermeasure Strategies
To inhibit action possibility	To inhibit action potentiality	1. To inhibit the power of the agent	1. To prevent the initial marshalling of the form of energy
		2. To reduce the power of the agent	2. To reduce the amount of energy marshalled
		3. To reduce the liability of the patient	8. To strengthen the structure which might be damaged by the energy transfer
	To eliminate and reduce action opportunity	4. To eliminate the opportunity for power- actualization (action) by separating the agent and the patient in space and time	5. To separate in space or time the energy being released from the susceptible structure
		5. To eliminate the opportunity for power-actualization (action) by separating the agent and the patient by interposition of a material barrier	6. To separate the energy being released from the susceptible structure by interposition of a material barrier
		6. To reduce the opportunity for power-actualization (action) by modifying the contact surface, subsurface, or basic structure which can be impacted	7. To modify the contact surface, subsurface, or basic structure which can be impacted
To prevent power-actualization		7. To prevent triggering conditions for power-actualization	3. To prevent the release of the energy
To control the actualized action		8. To modify power-actualization	4. To modify the rate of spatial distribution of release of energy from its source
		9. To control the state alterations (damage) of the patient	9. To move rapidly in detection and evaluation of damage and to counter its continuation and extension
		10. To stabilize and repair the state alterations (damage) of the patient	10. All those measures which fall between the emergency period following the damaging energy exchange and the final stabilization of the process

Table 2. Formalized description of countermeasures.

The formalized countermeasure types (corresponding to a classification of Action B in Figure 3) are described in the third column of Table 2. The countermeasure types are derived by adopting a negative view on causal action - eliminating, inhibiting and reducing its determinants shown in Figure 2, i.e. *inhibiting action possibility*, *preventing actualization*, *controlling the action*. Note that the sequence of countermeasures is different from the sequence of Haddon's original countermeasure strategies formulated on the basis of an energy transfer model (shown in the fourth column of Table 2). The numbers in column four indicate the original sequence. Preliminary results of the translation between the two approaches, based on two different kinds of models, are described in (Lind and Petersen, 2003).

A Critical Analysis of Barriers

The term “barrier” is widely used for analysis and assessment in in safety critical domains such as nuclear power plants. Despite its widespread use the term “barrier” is somewhat problematic and in the safety literature its specific meaning lacks consensus.

Below some of the problems associated with the use of the term “barrier” in a safety context are outlined. In order to clarify the term “barrier” a more precise terminology is proposed, based on an analysis of the semantic structure of countermeasures. In this analysis the model of causation described above has an important role to play.

The metaphorical misuse of the barrier concept

An important source of confusion is that the term “barrier” is used (or rather misused) metaphorically to denote things such as procedures and training that differ significantly from the original “material object” meaning of the term “barrier”. To understand these target objects as material objects is confusing and might lead to incorrect inferences.

Following Cognitive Semantics there is reason to believe that the barrier concept is grounded in our bodily experience of the physical world (Lakoff and Johnson, 1980). That is, the original meaning of the term “barrier” is most likely a material object that hinders physical motion through its location. Walls and fences are prototypical examples of barriers in this sense.

An early definition of the term “barrier” in a safety context is found in Haddon’s work on countermeasure (described above). Haddon’s sixth countermeasure strategy refers to the interposition of a “material barrier” as a way of separating the released energy and the susceptible structure (see Table 1). This specific meaning of the term “barrier” is close to its original “material object” meaning. Instead of hindering physical motion, the barrier hinders the flow of energy through its location.

It is important to note that the remaining nine countermeasure strategies are not realized by material barriers that separate the released energy from the susceptible structures. In fact, this is what distinguishes them from the sixth countermeasure strategy. The remaining countermeasure strategies refer to an entire range of different kinds of participating entities that are often referred to as “barriers” in the literature.

Using physically grounded concepts to describe abstract things is a common characteristic of everyday language. According to Lakoff and Johnson (1980) the metaphorical use of embodied concepts reflects a basic human condition for making sense and thinking about abstract things/concepts. However, in cases where the original meaning of a source concept provides a vague or even incorrect description of the target, metaphorical use of the source concept becomes problematic or even unacceptable. This is precisely why it is confusing the term barrier for things such as control systems, training and procedures.

Hollnagel's classification of different types of barrier systems includes barriers based on signs, so-called symbolic or immaterial barriers (Hollnagel, 1999). For instance, procedures can be perceived as symbolic barriers. Although Hollnagel is careful when defining these barrier types, including their dependency on the capability of intelligent agents to interpret signs, the metaphorical use of the term "barrier" remains a potential source of confusion. This is confirmed by Hale et al.'s comments on Hollnagel's extended barrier concept:

"...the behaviour which should be carried out and the circumstances under which that must take place; this is often defined in a written rule or procedure, but can be devised by an individual or group on the spot to meet unexpected circumstances: Hollnagel identifies these rule or procedures as 'immaterial barriers', but we see them only as elements in a total barrier function, since on their own they achieve nothing." Hale et al. (2004, p. 611).

Apart from the problems caused by the metaphorical misuse of the term "barrier", also the ontological status of "barrier" is ambiguous in the literature. For instance the MORT program refers to Haddon's countermeasure strategies as "barriers" (Johnson, 1973). As a result it is not always clear whether "barrier" is used to denote an action, an entity, or both.

Several attempts have been made to clarify the term "barrier". For instance, Svenson (1991) makes a distinction between so-called *barrier functions* (what is done to arrest the accident evolution) and *barrier systems* (things that maintain the barrier function). According to Svenson, barrier systems may be things such as an operator, an instruction, a physical separation, and an emergency control system. The means-end view on countermeasures adopted in this report, distinguishing between the countermeasure action and the participating entities, is similar to Svenson's distinction between functions and systems. In addition, however, the model of causation provides a systematic account of the relationship between countermeasures and the participating entities.

Towards a more precise terminology

Obviously it is not acceptable to have a vague conceptual basis for accident analysis and modeling of safety critical systems. The solution is not, however, to focus only on countermeasure performed by material objects. Although this would enable a well-defined use of the term "barrier", it fails to acknowledge other important types of countermeasures that go beyond those performed by (passive) material objects.

In order to clarify the term "barrier" an alternative terminology is proposed, based on an analysis of the causal structure of countermeasures prescribed by the model of causation described above. The model of causation gives primacy to the entities participating in action. These entities have well-defined roles such as e.g. agent and patient.

By means of applying the causal model at the level of countermeasure the relationship between countermeasures and the participating entities can be systematized. More specifically, the model can be used to make explicit the role of the entities participating in countermeasures. For instance, the agent refers to an entity with the power to perform the countermeasure. In the remaining part of the report the agent of the countermeasure is referred to as a *countermeasure agent* (see Figure 4).

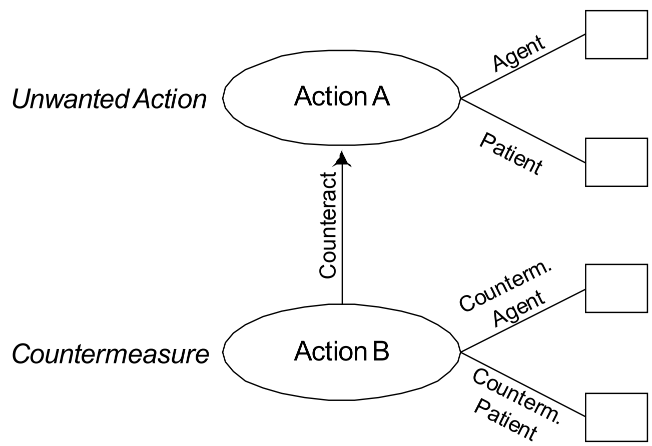

Figure 4. The means-end relation between an unwanted action and a countermeasure.

As opposed to the unwanted action (Action A), a countermeasure (Action B) is an intentional action performed to obtain a certain purpose, namely to oppose Action A. In the following advantage is taken of the fact that intentional actions have the same causal structure as simple physical actions.

It is proposed that the agent role of countermeasures can be used as a generalized "barrier" concept, covering the entire range of different entities with the power to perform a countermeasure. Apart from the fact that "countermeasure agent" is more general than the term "barrier" it does not have the drawbacks of the term "barrier" mentioned above. This means that the term "barrier" is restricted to denote material objects. Mentioned in passing the Swedish Nuclear Power Inspectorate has adopted a similar restricted view on the term "barrier" (SKI, 1998).

Although the concept of countermeasure agent solves some of the problems associated with the term "barrier" the present model of causation cannot account for the role that things such as procedures have in relation to countermeasures. It certainly does not make sense to model procedures as a countermeasure agent. After all it is not the procedure itself that is performing the countermeasure. A procedure is a *sign* (denoting something other then itself) that requires an intelligent agent with the capability of interpreting signs, in order to work (Hollnagel, 1999).

The role of signs cannot be accounted for by the current model of causation (or for that matter by Haddon's energy transfer model) because the agent is purely physical and incapable of interpreting signs. Clearly this is a limitation when we want to discuss the role of sign-based countermeasure in relation to the actions performed by intelligent agents, such as humans. In order to extend the analysis of countermeasures to include sign-based countermeasures, the causal model needs to be revised.

An extended view on countermeasures

The model of causation proposed above focuses on entities with the power and liability to take part in action. It is argued that an action can be opposed by 1) inhibiting action possibilities, 2) preventing the actualization of the action, or by 3) controlling the action once it is actualized. Under these headings a set of generic countermeasures where proposed (see Table 2).

In the following an attempt is made to widen the view on countermeasure to include also sign-based countermeasure, i.e. countermeasures performed as communicative actions. In order enable this a revised model of the unwanted phenomena, where an intelligent agent substitutes the causal agent, is proposed.

Just like a causal agent, an intelligent agent has the power to perform certain action in specific circumstances. However, when it comes to power-actualization there are fundamental differences between the two models. For instance, it is no longer correct to state that the power is actualized by specific extrinsic triggering condition. As opposed to a causal agent an intelligent agent has the capacity to make decisions. That is, the agent observes and interprets signs in its environment and depending on its preferences and goal(s) it chooses to act in a specific way (inside the envelope of possible actions).

The agent's *competence* is an additional intrinsic determinant of action. In order for the action to be possible it is not enough that the agent has the power to (physically) perform the action. In addition the agent typically needs to have a certain level of competence to perform the action. The revised model can be roughly described by the following sentence:

Competence & Potentiality & Opportunity & Decision = Actuality

Since the decision process depends on the agent's interpretation of signs it is possible, to some extent, to influence the decisions and hence the actions of the agent by shaping its environment. In order to account for the different ways in which the decision process can be influenced it is necessary to focus on *signs* and *sign production*.

The new type of countermeasures that emerges will have to be determined on the basis of an analysis of the role of signs and communicative action and how they can influence the actions of an agent (the interpreter).

Hollnagel (1999) has already argued that signs have a role to play in relation to countermeasures and facilitators. His notion of *symbolic barrier* refers to things that, by means of signifying something, can influence the actions of an intelligent agent (the interpreter). For instance, Hollnagel perceives a procedure as a symbolic barrier because it has the function to prescribe action. To prescribe is, however, not the only way a sign can influence an agent's action.

A comprehensive account of the different kinds of sign-based countermeasures should be based on the different ways in which the action of the intelligent agent (the interpreter) can, *in principle*, be counteracted by signs. For this purpose we will use Charles Morris' analysis of signs. His analysis of signs is developed from a behavioral perspective and he argues that signs have different modes of signifying, representing different ways of influencing the action/behavior of the interpreter. This corresponds to a so-called pragmatic view on signs (Morris, 1938).

A pragmatic view on signs

According to Morris (1946), a sign can be either *designative* (signifying characteristics of stimulus properties of objects), *appriasive* (signifying preferred objects or situations) or *prescriptive* (signifying required responses). This distinction between three modes of signifying is based on Mead's theory of action (Mead, 1938).

The different modes of signifying represent different ways in which the action/behavior of agents, that are capable of interpreting signs, can be influenced by signs.

A designative sign (a designator) signifies to the interpreter the characteristics (*discriminata*) of the situation (current and future). A designator does not signify the import of this object for its goals or what response is required toward the object in order to reach its goals. A designator influences what the agent *knows about* the current situations and its future developments. The agent makes judgments about the (possible) discrepancy between the current situation and the desired situation (according to his or her goals). If there is a discrepancy between the current situation and the desired situation there is a need to act in order to change the current situation. If there is no discrepancy the agent omits to act. Indirectly the designators influence this evaluation process since only what is known about the current and future situation can be compared to what is desired.

An appraisive sign (an appraisor) signifies to the interpreter a *preferential status* of an object in his or her behavior. In other words, it determines which objects rather than others the interpreter is disposed to favor in its behavior (his or her preferences). Hereby, the decision process is influenced.

A prescriptive sign (a prescriptor) signifies the required performance of a specific response to some situation. They are signs since they are substitutes in the control of behavior for the control some situation would exercise if it were present. A prescriptive sign may influence the action of the agent if there is a discrepancy between the prescribed action and action intended by the agent.

Means and ends of communicative actions

Apart from the possible modes of signification of signs, i.e. the different ways in which the interpreter in relation to his or her activity may interpret the sign, also the sign-production aspect needs to be discussed.

In order to obtain an instrumental understanding of how the behavior of an intelligent agent can be influenced it is important to perceive sign-production as an intentional act performed with a certain purpose, namely to influence the behavior of some intelligent agent in a specific way.

Speech act theory makes a distinction between different levels of action performed by the sign producer, corresponding to nested intentions of the sign producer. Austin (1962) identifies three distinct levels of action.

Locution: the act of saying something (meaning)
Illocution: what one does *in* saying it (illocutionary force)
Perlocution: what one does *by* saying it (achieving an effect).

Searle's (1969) notion of illocutionary force (assertive, directive, etc.) is similar to the different modes of signifying proposed by Morris. However, in opposition to Morris the illocutionary force of a sign refers to the intention of the sign producer and not the actual interpretation of the interpreter. In the remaining part of the report we use actually Morris' modes of signifying (designation, appraising and prescription) to describe illocutionary force.

Illocutionary acts can be described as <i,p> where i is the illocutionary force and p is the propositional content.

Since in most cases the primary intention is to bring about a perlocutionary effect, the illocution and the locution should be considered as (successive) means for this end. If, for example, a sign producer says "Shut the window" (locution) intending for the interpreter to understand this utterance as an order (illocution) and further intending that the interpreter should shut the window (perlocution).

The means-end structure of speech acts is shown in Figure 5. The perlocution is a causal action and its semantic structure can be accounted for by the causal model described in the previous section (by *agent* and a *patient* roles). The semantic structure of the illocution and the locution are different, however. The semantics of the locution is described by two roles; a *sign producer* and the *product*, i.e. the utterance produced, whereas two roles; *sign* and *interpreter* describe the semantics of the illocution. Since the primary purpose of the entity having a sign role is to mediate a message with a specific propositional content and illocutionary force it could be argued that the sign producer is implicit in the illocution. This is illustrated by the dotted line in Figure 5.

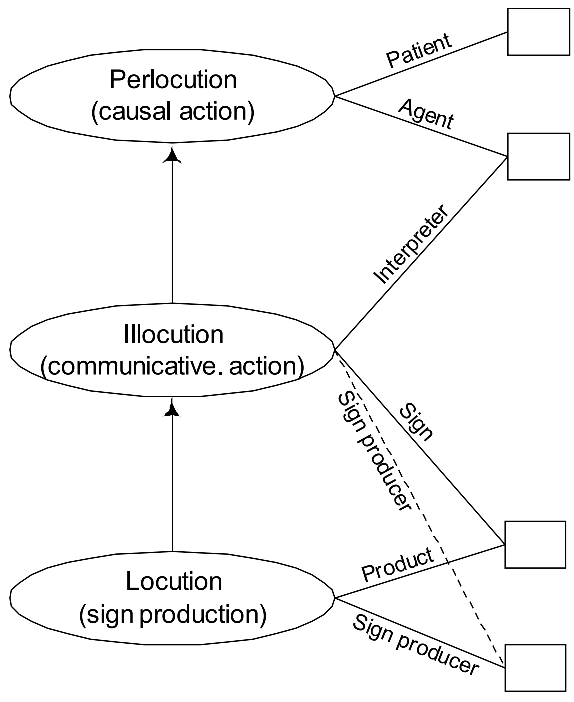

Figure 5. The means-end levels of speech acts.

Although the theory of speech acts is developed within a linguistic tradition, the theory covers any sort of sign usage across different media.

The function of procedures, for instance, can be explained in terms of Austin's levels of action. The writing the procedure corresponds to the *locution*, whereas the communication of the content of the procedure to the interpreter, typically several instructions, corresponds to the *illocution*. The *perlocution* is the intended effect of the procedure on the interpreter.

If the perlocution refers to an unwanted (hypothetical) action A the illocution is perceived as a countermeasure with the purpose to counteract that the perlocutionary agent performs A. Generally, the same perlocution can be implemented by different types of illocutions for which the propositional content refers to A more or less directly[1]. Below, three types of illocutionary countermeasures that complement the casual countermeasures described in the previous section are described.

[1] Consider another example where the perlocution itself is a countermeasure for another unwanted causal action. In this case the illocution should be perceived as a facilitator or a controlmeasure, i.e. an action with the intention to encouraging the interpreter to perform the countermeasure.

- By designating something, D it is possible to draw the attention of the agent to specific aspects of the situation and hereby prevent it from acting in a specific way (a way in which he would have acted, had D not been signified).
- By appraising something, A it is possible to influence the preferences of the agent and hereby prevent it from responding in a specific way (a way in which he would have acted had A not been signified).
- By prescribing something, P it is possible to prevent the agent from acting in a specific way (a way in which the agent would have acted had P not been signified).

  Example:
  Suppose that the perlocution corresponds to the action of not turning up the heat in the room, in which the agent is located (although it has turned cold). Below three different ways of bringing about the perlocution are described:

  1)
  Locution: p: "It is warmer in the adjacent room"
  Illocution: <designation, p>

  2)
  Locution: p: "It is cheaper to maintain a relatively low room temperature and wear much clothes to keep warm"
  Illocution: <appraising, p>

  3)
  Locution: p: "Do not turn up the heat"
  Illocution: <prescriptive, p>

Certain conditions must be satisfied for an illocution to be successful. An illocution is communicatively successful only if the interpreter recognizes the sign producer's illocutionary intention. This is what communication is about; anything more is more than just communication. Whether the illocution results in the desired perlocution depends on a several additional conditions. Even in cases where the interpreter understands the intention of the sign producer the perlocution does not necessarily follow. This is illustrated by means of the example above where the three cases indicate different levels of strengths between the illocutionary and the perlocutionary level of action.

In the third case there is straight forward relationship between the locution ("Do not turn up the heat"), the act of prescribing not to turn up the heat (the illocution) and the desired effect in the agent's behavior, i.e. "not to turn up the heat" (the perlocution). In the first case there is also a straight forward relationship between the locution ("It is warmer in the adjacent room") and the act of asserting that it is warmer in the adjacent room) (the illocution). Less direct is the connection between the locution and the desired effect with respect to the agent's behavior. Clearly there is no linguistic connection here for the signs does not mention turning up the heat. The indirect connection is inferential. The interpreter must infer what the sign-producer intends.

Actually this stresses the means-end levels of action referred to here. The perlocution reflects the (primary) intention with which the sign-producer is producing the sign. Whether the perlocution is actually performed is a matter of context that the sign-producer relies on the sign-user to rely on. This is true independent of the level of directness of the relation between the action levels.

Conclusions

This report has addressed a number of theoretical issues related to Haddon's countermeasure strategies.

- Based on a model of causation that generalizes Haddon's energy transfer model a formalization of Haddon's countermeasure strategies has been proposed. The formalized set of countermeasures has been derived by adopting a negative perspective on the determinants of causal action.
- Some of the problems associated with the term "barrier" have been outlined and an attempt has been made to clarify the use of the term based on the causal structure of countermeasures, i.e. the relationship between a specific countermeasure action and the participating entities.
- The scope of countermeasures was extended from physical actions to *communicative actions*, i.e. actions based on *signs* produced in order to influence the action of the agent interpreting them. Based on a discussion of the means-end levels of communicative action, using Speech Act Theory (Austin, 1962), a set of sign-based countermeasure types, complementing the causal countermeasure types, has been proposed.

References

Austin, J.L. (1962). How to do things with words. Harvard University Press. Cambridge, Massachusetts.

Brachman, R.J. (1979). On the Epistemological Status of Semantic Networks. In: Findler, N.V. (Ed.), Associative Networks: Representation and Use of Knowledge by Computers. New York: Academic Press.

Fillmore, C.J. (1968). The case for case. In: Batch, E. and Harms, R.T. (Eds.) Universals in Liguistic Theory. Holt, Reinhart and Winston, New York.

Haddon, W. (1973). Energy Damage and the Ten Countermeasure Strategies. Human Factors, 15(4), 355-366.

Hale, A. et al. (2004). Managing safety barriers and controls at the workplace. PSAM7/ESREL'04 International Conference on Probabilistic Safety Assessment and Managment, 608-613, June 14-18, 2004, Berlin, Germany.

Harré, R. and Madden, E.H. (1975). Causal Powers: A theory of natural necessity. Basil Blackwell, Oxford.

Hollnagel, E. (1999). Accidents and Barriers. Proceedings of the 7th Conference on Cognitive Science Approaches to Process Control (CSAPC '99), Villeneuve d'Asc (FR), 21-24 Sept 1999. Hoc, J.M.; Millot, P.; Hollnagel; E. and Cacciabue, P.C. (eds.), (1999), p. 175-180.

Hume, D. (2000). A Treatise of Human Nature. Oxford University Press.

Johnson, W.G. (1973). MORT – The Management Oversight and Risk Tree. SAN 821-2. US Atomic Energy Commission.

Kingston, J. (2002). 3CA Control Change Cause Analysis Manual. The Noordwijk Risk Initiative Foundation.

Kingston, J., Nertney, R., Frei, R. and Schallier, P. (2004). Barrier Analysis Analysed in MORT Perspective. PSAM7/ESREL'04 International Conference on Probabilistic Safety Assessment and Managment, 364-369, June 14-18, 2004, Berlin, Germany.

Lakoff and Johnson (1980). Metaphors We Live By. The University of Chicago Press.
Morris, C. (1946). Signs, Language and Behavior. Prentice Hall, New York.

Lind, M. (2000). Possibilities for Action. Report from Center for Human-Machine Interaction. CHMI-7-2000.

Lind, M. and Petersen, J. (2003). A Review of Barrier Concepts. In: NKS project NKS-R-07: Barriers, Control and Management – Report from the pilot phase.

Mead, G.H. (1938). Philosophy of the act. The University of Chicago Press.

Morris, C. (1938). Foundations of the Theory of Signs. International Encyclopedia of Unified Science, Vol. 1, No. 2. Chicago: University of Chicago Press.

Morris, C. (1946). Signs, Language and Behavior. Prentice Hall, New York.

Rasmussen, J. (1991). Event Analysis and the Problem of Causality. In: Rasmussen, J. et al. (Eds.), Distributed Decision Making: Cognitive Models for Cooperative Work. John Wiley & Sons Ltd.

SKI (1998). Swedish Nuclear Power Inspectorate's Regulations Concerning Safety in Certain Nuclear Facilities. SKIFS 1998:1

Sowa, J.F. (1984). Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley.

Svenson, O. (1991). The Accident Evolution and Barrier Function Model Applied to Incident Analysis in the Processing Industries. Risk Analysis, Vol. 11, No. 3.

Searle, J.R. (1969) Speech Acts. Cambridge University Press.

Title	Countermeasures and Barriers
Author(s)	Johannes Petersen
Affiliation(s)	Ørsted DTU, Automation, Technical University of Denmark
ISBN	87-7893-174-6 *(Electronic report)*
Date	October 2005
Project	NKS_R_2002_07
No. of pages	19
No. of tables	2
No. of illustrations	5
No. of references	22

Abstract

In 1973 Haddon proposed ten strategies for reducing and avoiding damages based on a model of potential harmful energy transfer (Haddon, 1973). The strategies apply to a large variety of unwanted phenomena. Haddon's pioneering work on countermeasures has had a major influence on later thinking about safety. Considering its impact it is remarkable that the literature offers almost no discussions related to the theoretical foundations of Haddon's countermeasure strategies. The present report addresses a number of theoretical issues related to Haddon's countermeasure strategies, which are:

- A reformulation and formalization of Haddon's countermeasure strategies
- An identification and description of some of the problems associated with the term "barrier"
- Suggestions for a more precise terminology based on the causal structure of countermeasures
- Extending the scope of countermeasures to include sign-based countermeasures

Key words: Countermeasures, barriers, Haddon's strategy, theoretical foundation

Available on request from the NKS Secretariat, P.O.Box 49, DK-4000 Roskilde, Denmark.
Phone (+45) 4677 4045, fax (+45) 4677 4046, e-mail nks@nks.org, www.nks.org